08-17-00

A

UTILITY PATENT APPLICATION TRANSMITTAL
(Only for new nonprovisional applications under 37 CFR 1.53(b))

Docket No.	:	39981/DAC/X2
Inventor(s)	:	Rex A. Hill
Title	:	DISTRIBUTED SOURCE LEARNING FOR DATA COMMUNICATION SWITCH
Express Mail Label No.	:	EL 496225091US

ADDRESS TO: Assistant Commissioner for Patents
Box Patent Application
Washington, D.C. 20231

Date: August 16, 2000

1. X **FEE TRANSMITTAL FORM** *(Submit an original, and a duplicate for fee processing).*

2. **IF A CONTINUING APPLICATION**

____ This application is a of patent application No. .

Prior application information: Examiner ; Group Art Unit:

____ This application claims priority pursuant to 35 U.S.C. §119(e) and 37 CFR §1.78(a)(4), to provisional Application No. .

3. **APPLICATION COMPRISED OF**

Specification

28 Specification, claims and Abstract (total pages)

Drawings

7 Sheets of drawing(s) (FIGS. 1 to 7)

Declaration and Power of Attorney

____ Newly executed
X Unexecuted declaration
____ Copy from a prior application (37 CFR 1.63(d))(for continuation and divisional)

4. ____ **Microfiche Computer Program** *(Appendix)*

5. ____ **Nucleotide and/or Amino Acid Sequence Submission** *(if applicable, all necessary)*

____ Computer Readable Copy
____ Paper Copy (identical to computer copy)
____ Statement verifying identity of above copies

6. **ALSO ENCLOSED ARE**

____ Preliminary Amendment
____ A Petition for Extension of Time for the parent application and the required fee are enclosed as separate papers
____ Small Entity Statement(s)
____ Statement filed in parent application, status still proper and desired

UTILITY PATENT APPLICATION TRANSMITTAL
(Only for new nonprovisional applications under 37 CFR 1.53(b))

Docket No.: 39981/DAC/X2

- ___ Copy of Statement filed in provisional application, status still proper and desired
- ____ An Assignment of the invention with the Recordation Cover Sheet and the recordation fee are enclosed as separate papers
- ____ This application is owned by pursuant to an Assignment recorded at Reel , Frame
- ____ Information Disclosure Statement (IDS)/PTO-1449
  - ____ Copies of IDS Citations
- ____ Certified copy of Priority Document(s) *(if foreign priority is claimed)*
- ____ English Translation Document *(if applicable)*
- X Return Receipt Postcard (MPEP 503) (should be specifically itemized).
- ____ Other

7. CORRESPONDENCE ADDRESS

CHRISTIE, PARKER & HALE, LLP, P.O. BOX 7068, PASADENA, CA 91109-7068

Respectfully submitted,

CHRISTIE, PARKER & HALE, LLP

By Jun-Young E. Jeon

Jun-Young Jeon
Reg. No. 43,693
626/795-9900

JEJ/dp

FEE TRANSMITTAL
UTILITY PATENT APPLICATION

DATE: August 16, 2000

Docket No.	:	39981/DAC/X2
Inventor(s)	:	Rex A. Hill
Title	:	DISTRIBUTED SOURCE LEARNING FOR DATA COMMUNICATION SWITCH

FEE CALCULATIONS					
	CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE	CALCULATIONS
A	TOTAL CLAIMS	64 - 20 =	44	44 x $9.00	$396.00
B	INDEPENDENT CLAIMS	3 - 3 =	0	0 x $39.00	$
C	SUBTOTAL	SMALL ENTITY FEE = A + B LARGE ENTITY FEE = 2 X (A + B)			792.00
D	BASIC FEE	SMALL ENTITY FEE = $345.00 LARGE ENTITY FEE = $690.00			690.00
E	MULTIPLE-DEPENDENT CLAIMS FEE	SMALL ENTITY FEE = $130.00 LARGE ENTITY FEE = $260.00			
F	TOTAL FILING FEE (ADD LINES C, D, AND E)				$1,482.00
List Independent Claims: 1, 4 and 61					

METHOD OF PAYMENT

_____ Payment Enclosed:

_____ The Commissioner is hereby authorized to charge any fees under 37 CFR 1.16 and 1.17 which may be required during the **entire pendency** of the application to Deposit Account No. 03-1728. Please show our docket number with any charge or credit to our Deposit Account. **A duplicate copy of this sheet is enclosed.**

Respectfully submitted,

CHRISTIE, PARKER & HALE, LLP

By Jun-Young E. Jeon

Jun-Young Jeon
Reg. No. 43,693
626/795-9900

JEJ/dp
JEJ PAS268100 1-*-8/16/00 11 44 AM

DISTRIBUTED SOURCE LEARNING FOR DATA COMMUNICATION SWITCH

FIELD OF THE INVENTION

The present invention relates to devices for source learning and, more particularly, to devices for distributed source learning in a data communication switch.

BACKGROUND OF THE INVENTION

Data communication switches interconnect network devices residing in different network domains. Such switches typically include a plurality of switching modules for switching data traffic between external network devices and a centralized management module for configuring the switching modules. Part of the switching module configuration is "source learning." Source learning is the process of dynamically learning associations between ports and the addresses of network devices they support by reviewing source addresses in inbound packets. By making such address-port associations, packets can be advantageously forwarded only on the ports of the switch supporting packet destinations rather than being "flooded" on all ports.

In a conventional source learning process, source addresses in packets are reviewed by a switching module upon ingress and unknown source addresses are submitted to the source learning function resident on a centralized management module for processing. The management module configures the address-port association on the

switching modules such that future inbound packets destined to that address can be forwarded without unnecessary flooding.

While the source learning process has resulted in bandwidth savings in the form of reduced flooding, such savings have come at a price. Reliance on a centralized management entity for source learning has required a special internal protocol for flagging packets requiring source learning for capture by the management entity and has caused bottlenecks at the management module when many packets requiring source learning arrive at different switching modules within a short period of time.

SUMMARY OF THE INVENTION

The invention provides an efficient method and apparatus for accomplishing source learning in a data switch of the type having a plurality of switching modules, each supporting one or more external network devices and a backplane interconnecting the switching modules. Each switching module has logic resident thereon for performing distributed source learning, including configuring unknown source addresses “seen” in inbound packets and for making available or notifying the other switching modules that such source addresses were “seen” on a port thereof. Address-port associations are thereby configured on the switch using distributed logic, i.e. without intervention by a centralized management entity. Packets having unknown source addresses are replicated at the first switching module to enable packet forwarding across the backplane to proceed in parallel with source learning. Exchange of source learning information between switching modules is made “out of band” on a bus interconnecting the switching modules. Packets having unknown destination addresses are replicated at the first

switching module where one copy is sent to a multicast queue for transmission and another copy is sent to a software module to find the destination address. Once the destination address is found, the multicast flow is interrupted, data is buffered for a period of time to ensure flow integrity (by keeping packets in correct order), and then the data flow is continued to a unicast queue for transmission.

These and other aspects of the invention can be understood by reference to the following detailed description, taken in conjunction with the accompanying drawings, which are briefly described below.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a diagram of an exemplary data communication switch;

Figure 2 is a diagram of an exemplary switching module within the data communication switch of Figure 1;

Figure 3 is a flow diagram illustrating a source and destination indexing protocol for a known source and known destination according to Figure 2;

Figure 4 is a flow diagram illustrating a source and destination indexing protocol for an unknown source and known destination according to Figure 2;

Figure 5 is a flow diagram illustrating a source and destination indexing protocol for an unknown source and unknown destination and known source and unknown destination according to Figure 2;

Figure 6 is a flow diagram illustrating the source learning protocol acquiring the destination address according to Figure 2; and

Figure 7 is a flow diagram illustrating the source learning protocol associating the source address with a port.

DETAILED DESCRIPTION

Figure 1 illustrates switch 100 in one embodiment of the present invention. Switch 100 includes switching backplane 102 driven by switching modules 104, 106, and 108, where switching module 104 is coupled to host device 110, switching module 106 is coupled to host device 112, and switching module 108 is coupled to host device 114. Additionally, switching modules 104 and 106 are coupled to each other by control path 116 and switching modules 106 and 108 are coupled to each other by control path 118. Each module104, 106, and 108 interfaces with backplane 102 over data path 120, 122, and 124, respectively, to transmit packet data to backplane 102 and receive packet data from the backplane 102. For example, host device 110 preferably determines whether a destination device is on the same IP or IPX network by comparing its layer 3 addresses to the destination layer 3 address. If the destination address comparison indicates that the destination device is on the same network, an Address Resolution Protocol (ARP) message is sent from host 110 to retrieve the layer 2 address of the destination, and bridging is used to transmit the packet data. If the destination device is not on the same network, an ARP message is sent to retrieve the layer 2 default Media Access Control (MAC) address of the first router which will lead to the destination device, and routing is used to transmit the packet data. In the latter case, while the layer 2 default MAC address constantly changes to reflect the next router address leading to the destination device, the

layer 3 IP destination address preferably stays constant to reflect where the packet is going.

Figure 2 is a block diagram of switching module 200, which may be similar to switching module 104 of FIG. 1. Switching module 200 preferably has a source learning capability, which will be described in reference to Figure 2. Module 200 includes access controller 202 coupled to switching controller 204. Access controller 202 receives packets from host devices, operates on them, and transmits them to switching controller 204. Access controller 202 also receives packets from switching controller 204, operates on them, and transmits them to host devices. Switching controller 204 is not only coupled to access controller 202 but is coupled to queue controller 206 as well. Switching controller 204, similar to access controller 202, receives packets from access controller 202, processes them, and transmits them to queue controller 206. Switching controller 204 also receives packets from queue controller 206, processes them, and transmits them to access controller 202. Queue controller 206 includes unicast packet buffer (UPB) 218, multicast packet buffer (MPB) 220 and lock table 222. Queue controller 206 is coupled to many elements, including source address resolution element (SARE) 208, destination address resolution element (DARE) 210, unicast queue 212, multicast queue 214, queue identification (QID) 216, and source learning element 224 (where source learning element 224 is coupled to software table 226 and pseudoCAM (PCAM) 228, which may be implemented in hardware, software, or both).

Queue controller 206 preferably receives a data packet from switching controller 204, SARE 208 determines whether the source address is known for the packet, DARE 210 determines whether the destination address is known for the packet, and QID 216

assigns a port number, priority, and bandwidth to the packet. Then, queue controller 206 stores the packet in unicast queue 212 or multicast queue 214 to be transmitted when its priority for the particular port is reached.

Source and Destination Conditions:

One embodiment of the present invention is a novel source learning technique using multiple switching modules coupled together on a single backplane. The single backplane architecture preferably allows for source learning and transmitting determinations to be made for packets having different source and destination conditions in a single path. Packets with a known source address and a known destination address preferably are not sent to a source learning element and are transmitted to a unicast queue for transmission and forwarding. Packets with an unknown source address and a known destination address preferably are sent to the source learning element for source learning and concurrently transmitted to a unicast queue for transmission and forwarding. Packets with a known source address and an unknown destination address preferably are not sent to the source learning element and are transmitted to a multicast queue for transmission and forwarding. Packets with an unknown source and an unknown destination preferably are sent to the source learning element for source learning and concurrently transmitted to a multicast queue for transmission and forwarding.

Therefore, the source learning technique in this embodiment preferably processes the following four categories of packets in a single flow path: (1) known source address and known destination address; (2) unknown source address and known destination address; (3) known source address and unknown destination address; and (4) unknown source address and unknown destination address. In the case where the destination

address is known, flow integrity typically is not an issue since the unicast queue normally is the only queue being used. In other embodiments, the source learning technique may use more than one flow path to process the four categories of packets.

Known Source and Known Destination:

Referring to Figure 3, a packet is received at queue controller 206 (302) from switching controller 204. Upon receiving the packet, a lookup operation is performed to determine the source address in SARE 208 (304). If the source address is not found (306), the packet is tagged for source learning in source learning element 224 (308). If the source address is found (306), a lookup operation is performed to determine the destination address in DARE 210 (310). If the destination address is not found (312), the packet is defined for flooding (314) and source learning element 224 is notified (316) so that it can search for the destination address in the switching modules. If the destination address is found (312), as is the case here, the packet may follow one of the following three paths depending on the state of the destination address: last multicast packet path, first unicast packet path, and neither last multicast nor first unicast packet path. These paths ensure flow integrity for the packets. As a result, all multicast packets preferably are transmitted to this destination address before any unicast packets are transmitted there.

Last Multicast Packet Path:

Once the packet is found to have both its source and destination addresses associated with a port, queue controller 206 preferably performs a check to see if the flow state is marked as the last multicast packet (318). If the flow state is marked as the last multicast packet, thus indicating that the packet is defined for flooding (320), a lock bit is

set in multicast packet buffer (MPB) 220 (322) internal to queue controller 206, and the flow state is changed from "last multicast packet" to "first unicast packet" (324). In such state, the packet is still flooded. Thus, referring to Figure 5, QID 216 preferably defines the port for flooding and subsequently finds the priority and bandwidth to be applied to the packet (502). Additionally, to ensure that all multicast packets are transmitted to this destination address before any unicast packets are transmitted there, a lock bit in lock table 222 internal to queue controller 206 is set (504) when the lock bit is set in MPB 220 (506). If the lock bit is not set in MPB 220, the lock bit is not set in lock table 222 (506). In either case, the packet is thereafter stored in multicast queue 214 until the bandwidth is available at the specified priority for the port (508). Once bandwidth is available at the specified priority for the port, the packet is transmitted (510), the lock bit is cleared in lock table 222 (510) and the packet is tested to see if the source address is known (512). In this case, nothing more is done since the source address is known (514). If the source address were unknown, the packet would be sent to source learning element 224 (516).

First Unicast Packet Path:

Referring to Figure 3, if the flow state is not marked as the last multicast packet (318), a test is performed to see if the flow state is marked as the first unicast packet (326). If, as in this case, the flow state is marked as the first unicast packet, a lock bit is set in a unicast packet buffer (UPB) 218 (328). Next, QID 216 preferably defines the port, priority, and bandwidth for the packet (330) and the packet is stored in unicast queue 212 until bandwidth is available at the specified priority for the port (332). Once bandwidth is available at the specified priority for the port, a check preferably is performed to see if the lock bit in UPB 218 is clear (334). If the lock bit in UPB 218 is clear, the packet is

transmitted (336). As is the case here, the lock bit in the UPB 218 is set, and consequently, a test is performed to see if the lock bit in lock table 222 is clear (338). If the lock bit in lock table 222 is clear, the packet is transmitted (336). If the lock bit in lock table 222 is not clear, the packet is buffered until it is cleared by the transmission of the last multicast packet (340). Once the last multicast packet has been transmitted, then this packet is transmitted (336).

Neither Last Multicast nor First Unicast Packet Path:

Referring to Figure 3, if the flow state is not marked as a last multicast packet or first unicast packet, the packet is forwarded to QID 216 so that the port, priority, and bandwidth can be defined for the packet (330) and the packet will be stored in unicast queue 212 until bandwidth is available at the specified priority for the port (332). Once bandwidth is available at the specified priority for the port, a check preferably is performed to see if the lock bit in UPB 218 is clear (334). If the lock bit in UPB 218 is clear, as is the case here, the packet is transmitted (336). If the lock bit in the UPB 218 is set, a test is performed to see if the lock bit in lock table 222 is clear (338). If the lock bit in lock table 222 is clear, the packet is transmitted (336). If the lock bit in lock table 222 is not clear, the packet is buffered until the lock bit in lock table 222 is cleared by the transmission of the last multicast packet (340). Once the last multicast packet has been transmitted, then this packet is transmitted (336).

<u>Unknown Source and Known Destination:</u>

Referring to Figure 3, a packet is received at queue controller 206 (302) from switching controller 204. Upon receiving the packet, a lookup operation is performed to determine the source address in SARE 208 (304). If the source address is found (306), a

lookup operation is performed to find the destination address (310). If the source address is not found (306), as is the case here, the packet is tagged for source learning in source learning element 224 (308). Referring to Figure 4, after the packet is tagged for source learning (308), it is further processed so that a destination lookup (402) can be performed in DARE 210. If the destination address is not found (404), the packet is defined for flooding (406) and source learning 224 is notified so that it can search for the destination address (408). If the destination address is found, which is the case here, the packet may follow one of the following three paths depending on the state of the destination address: last multicast packet path, first unicast packet path, and neither last multicast nor first unicast packet path. These paths ensure flow integrity for the packets. As a result, all multicast packets preferably are transmitted to this destination address before any unicast packets.

Last Multicast Packet Path:

Once the packet is found to have an unknown source address and a known destination address, queue controller 206 preferably performs a check to see if the flow state is marked as the last multicast packet (410). If the flow state is marked as the last multicast packet, thus indicating that the packet is defined for flooding (412), a lock bit is set in MPB 220 (414) internal to queue controller 206 and the flow state is changed from last multicast packet to first unicast packet (416). In such state, the packet is still flooded. Thus, referring to Figure 5, QID 216 preferably defines the port for flooding and subsequently finds the priority and bandwidth to be applied to the packet (502). Additionally, to ensure that all multicast packets are transmitted to this destination address before any unicast packets are transmitted to this destination address, a lock bit in

lock table 222 internal to queue controller 206 is set (504) when the lock bit is set in MPB 220 (506). If the lock bit is not set in MPB 220, the lock bit is not set in lock table 222 (506). In either case, the packet is thereafter stored in multicast queue 214 until bandwidth is available at the specified priority for the port (508). Once bandwidth is available at the specified priority for the port, the packet is transmitted (510), the lock bit is cleared in lock table 222 (510) and the packet is tested to see if the source address is known (512). In this case, the source address is unknown and the packet is sent to source learning 224 (516) so that its source address can be associated with its particular port. Referring to Figure 7, a request is received by source learning 224 to learn a source address (702). Upon processing the request, a layer 2 source MAC which relates to a port is stored in a software table 226 (704). This software table 226 may be used for many things, for example, source learning 224 may use it to inform its own and/or other switching modules of a new source address and/or source learning 224 may use it to allow access for its own and/or other modules to read and/or write to the software table 226. Thereafter, source learning software in source learning element 224 will place the source MAC in a hardware table pseudo CAM 228 (706) and then will wait for another request to perform source learning. If the source address were known, the packet would not be sent to source learning element 224 (514) and nothing more would be done.

First Unicast Packet Path:

Referring to Figure 4, if the flow state is not marked as the last multicast packet (410), a test is performed to see if the flow state is marked as the first unicast packet (418). If, as in this case, the flow state is marked as the first unicast packet, the lock bit is set in UPB 218 (420). Next, QID 216 defines the port, priority, and bandwidth for the

packet (422) and the packet will be stored in unicast queue 212 until bandwidth is available at the specified priority for the port (424). Once bandwidth is available at the specified priority for the port, a check preferably is performed to see if the lock bit in UPB 218 is clear (426). If the lock bit in UPB 218 is clear, the packet is transmitted (428) and sent to source learning element 224 (430). As is the case here, the lock bit in the UPB 218 is set, and consequently, a test is performed to see if the lock bit in lock table 222 is clear (432). If the lock bit in lock table 222 is clear, the packet is transmitted (428) and sent to source learning 224 (430). If the lock bit in lock table 222 is not clear, the packet is buffered until the lock bit in lock table 222 is cleared by the transmission of the last multicast packet (434). Once the last multicast packet has been transmitted, then this packet is transmitted (428) and sent to source learning element 224 (430). The packet is sent to source learning 224 (700) so that its source address can be associated with its particular port. Referring to Figure 7, a request is received by source learning 224 to learn a source address (702). Upon processing the request, a layer 2 source MAC which relates to a port is stored in a software table 226 (704). This software table 226 may be used for many things, for example, source learning 224 may use it to inform its own and/or other switching modules of a new source address and/or source learning 224 may use it to allow access for its own and/or other modules to read and/or write to the software table 226. Thereafter, source learning software in source learning element 224 will place the source MAC in a hardware table pseudo CAM 228 (706) and then will wait for another request to perform source learning.

Neither Last Multicast nor First Unicast Packet Path:

Referring to Figure 4, if the flow state is not marked as a last multicast packet or first unicast packet, the packet preferably is forwarded to QID 216 so that the port, priority, and bandwidth can be defined for the packet (422) and the packet is stored in unicast queue 212 until bandwidth is available at the specified priority for the port (424). Once bandwidth is available at the specified priority for the port, a check preferably is performed to see if the lock bit in UPB 218 is clear (426). If the lock bit in UPB 218 is clear, as is the case here, the packet is transmitted (428) and sent to source learning element 224 (430). If the lock bit in the UPB 218 is set, a test is performed to see if the lock bit in lock table 222 is clear (432). If the lock bit in lock table 222 is clear, the packet is transmitted (428) and sent to source learning element 224 (430). If the lock bit in lock table 222 is not clear, the packet is buffered until the lock bit in lock table 222 is cleared by the transmission of the last multicast packet (434). Once the last multicast packet has been transmitted, then this packet is transmitted (428) and sent to source learning element 224 (430). The packet is sent to source learning 224 (700) so that its source address can be associated with its particular port. Referring to Figure 7, a request is received by source learning 224 to learn a source address (702). Upon processing the request, a layer 2 source MAC which relates to a port is stored in a software table 226 (704). This software table 226 may be used for many things, for example, source learning 224 may use it to inform its own and/or other switching modules of a new source address and/or source learning 224 may use it to allow access for its own and/or other modules to read and/or write to the software table 226. Thereafter, source learning software in source learning element 224 will place the source MAC in a hardware table pseudo CAM 228 (706) and then will wait for another request to perform source learning.

Known Source and Unknown Destination:

Referring to Figure 3, a packet is received at queue controller 206 (302). Upon receiving the packet, a lookup operation is performed to determine the source address 208 (304). If the source address is not found (306), the packet is tagged for source learning in source learning element 224 (308). If the source address is found (306), as is the case here, a lookup operation is performed to determine the destination address in DARE 210 (310). If the destination address is found (312), a check is performed to see if the flow state is marked as a last multicast packet (318). If the destination address is not found (312), as is the case here, the packet is defined for flooding (314) and source learning element 224 is notified so that it can search for the destination address in the switching modules (316). Referring to Figure 6, source learning element 224 receives a request to find the destination address (602). Once the request has been received, source learning element 224 uses its software to look in its own modules and others to find the destination address (604). If the destination address is not found (606), the flood of packets are allowed to continue (608), a request to find the destination is again received (602), the software is used to search for the destination address (604), and a test is performed to see if the destination address was found this time (606). This process preferably continues until the destination address is found. If the destination address is found, QID 216 defines a port, priority, and bandwidth for the packet (610), an entry is created in PCAM 228 for the new destination address (612), and the flow state is set to last multicast (614) so that the last remaining packet is transmitted to this destination over multicast queue 214 before the first unicast packet is transmitted to this destination over unicast queue 2122.

Referring to Figure 5, after the packet is defined for flooding (314) and source learning element 224 is notified so that it can search for the destination address in the switching modules (316), QID 216 defines the port for flooding and subsequently finds the priority and bandwidth to be applied to the packet (502). Additionally, to ensure that all multicast packets are transmitted to this destination address before any unicast packets are transmitted there, a lock bit in lock table 222 internal to queue controller 206 is set (504) when the lock bit is set in MPB 220 (506). If the lock bit is not set in MPB 220, then the lock bit in lock table 222 is not set (506). In either case, the packet is thereafter stored in a multicast queue 214 until bandwidth is available at the specified priority for the port (508). Once bandwidth is available at the specified priority for the port, the packet is transmitted (510), the lock bit is cleared in lock table 222 (510), and the packet is tested to see if the source address is known (512). In this case, the source address is known and nothing more is done (514). If the source address is unknown, the packet is sent to source learning element 224 (516) so that the source address can be associated with its particular port.

<u>Unknown Source and Unknown Destination:</u>

Referring to Figure 3, a packet is received at queue controller 206 (302). Upon receiving the packet, a lookup operation is performed to determine the source address 208 (304). If the source address is found (306), a lookup operation is performed to determine the destination address (310). If the source address is not found (306), as is the case here, the packet is tagged for source learning in source learning element 224 (308). Referring to Figure 4, after the packet is tagged for source learning (308), it is further processed so that a destination lookup (402) can be performed in DARE 210. If the destination is

found, the packet may follow one of three paths, depending on the state of the destination address. If the destination address is not found (404), as is the case here, the packet is defined for flooding (406) and source learning 224 is notified so that it can search for the destination address (408). Referring to Figure 6, source learning element 224 receives a request to find the destination address (602). Once received, source learning element 224 uses its software to look in its own modules and others to find the destination address (604). If the destination address is not found (606), the flooding of packets is allowed to continue (608), a request to find the destination is again received (602), the software is used to search for the destination address (604), and a test is performed to see if the destination address was found this time (606). This process preferably continues until the destination address is found. If the destination address is found, QID 216 defines a port, priority, and bandwidth for the packet (610), an entry is created in PCAM 228 for the new destination address (612), and the flow state is set to last multicast (614) so that the last remaining packet is transmitted over multicast queue 214 before the first unicast packet is transmitted over unicast queue 212. Referring to Figure 5, after the packet is defined for flooding (314) and source learning element 224 is notified so that it can search for the destination address in the switching modules (316), QID 216 will define the port for flood and will subsequently find the priority and bandwidth to be applied to the packet (502). Additionally, to make sure all multicast packets are transmitted to this destination address before any unicast packets are transmitted to this destination address, a lock bit in lock table 222 internal to queue controller 206 is set (504) when the lock bit is set in MPB 220 (506). If the lock bit is not set in MPB 220, then the lock bit in lock table 222 is not set (506). In either case, the packet is thereafter stored in a multicast queue 214 until

bandwidth is available at the specified priority for the port (508). Once bandwidth is available at the specified priority for the port, the packet is transmitted (510), the lock bit is cleared in lock table 222 (510), and the packet is tested to see if the source address is known (512). In this case, the source address is unknown and the packet is sent to source learning 224 (516) so that its source address can be associated with its particular port. Referring to Figure 7, a request is received by source learning 224 to learn a source address (702). Upon processing the request, a layer 2 source MAC which relates to a port is stored in a software table 226 (704). This software table 226 may be used for many things, for example, source learning 224 may use it to inform its own and/or other switching modules of a new source address and/or source learning 224 may use it to allow access for its own and/or other modules to read and/or write to the software table 226. Thereafter, source learning software in source learning element 224 will place the source MAC in a hardware table pseudo CAM 228 (706) and then will wait for another request to perform source learning. If the source address were known, the packet would not be sent to source learning element 224 (514) and nothing more would be done.

It will be appreciated by those of ordinary skill in the art that the invention can be embodied in other specific forms without departing from the spirit or essential character hereof. The present description is therefore considered in all respects to be illustrative and not restrictive. The scope of the invention is indicated by the appended claims, and all changes that come within the meaning and range of equivalents thereof are intended to be embraced therein.

I claim:

1. A switching module for a data communication switch having a plurality of switching modules interconnected over a backplane, comprising:

a first port for receiving packet data including a source address and a destination address;

means for checking whether a source association between the source address and the first port has been made; and

means for making the source association and providing source association information to other switching modules,

wherein the means for making the source association makes the source association when the source association has not been made.

2. The switching module of claim 1, further comprising:

means for checking whether a destination association between the destination address and a second port has been made; and

means for making the association while current packets are flooded over a multicast queue.

3. The switching module of claim 2, wherein when the destination association has been made, packets are transitioned from traversing the multicast queue to a unicast queue.

4. A method for communicating data over a switch having a plurality of switching modules interconnected over a backplane, each switching module having a capability to perform source learning, the method comprising the steps of:

receiving a packet on a first port, the packet including a source address and a destination address;

determining whether the source address is associated with the first port;

determining whether the destination address is associated with a second port;

determining a flow state;

storing the packet; and

transmitting the packet when a specified bandwidth is available at a specified priority for a specified port.

5. The method of claim 4 further comprising the step of: specifying the port, the priority and the bandwidth for the packet if the flow state is neither "last multicast packet" nor "first unicast packet."

6. The method of claim 5, wherein the step of storing the packet comprises the step of storing the packet in a unicast queue until the specified bandwidth is available at the specified priority for the specified port.

7. The method of claim 6 further comprising the step of checking to ensure that a lock bit in a unicast packet buffer has been cleared when the specified bandwidth is available at the specified priority for the specified port.

8. The method of claim 7 further comprising the step of performing a check to ensure that a lock bit in a lock table has been cleared before the packet is transmitted if the lock bit in the unicast packet buffer has not been cleared.

9. The method of claim 4 further comprising, if the flow state is "last multicast packet", the steps of: defining the packet for flooding; setting a lock bit in a multicast packet buffer; and changing the flow state to "first unicast packet."

10. The method of claim 9 further comprising the step of specifying the port, the priority and the bandwidth for the packet.

11. The method of claim 10 further comprising the step of setting a lock bit in a lock table if the lock bit is set in the multicast packet buffer.

12. The method of claim 11, wherein the step of storing the packet comprises the step of storing the packet in a multicast queue until the specified bandwidth is available at the specified priority for the specified port.

13. The method of claim 12 further comprising the step of clearing the lock bit in the lock table when the packet is transmitted.

14. The method of claim 13 further comprising the step of sending the packet to a source learning element if the source address is not associated with the first port.

15. The method of claim 14 further comprising the step of source learning in which the source address is associated with the first port and association information is stored for inquiries or transmittal.

16. The method of claim 4 further comprising the step of setting a lock bit in a unicast packet buffer if the flow state is "first unicast packet."

17. The method of claim 16 further comprising the step of specifying the port, the priority, and the bandwidth for the packet.

18. The method of claim 17, wherein the step of storing the packet comprises the step of storing the packet in a unicast queue until the specified bandwidth is available at the specified priority for the specified port.

19. The method of claim 18 further comprising the step of performing a check to ensure that the lock bit in the unicast packet buffer has been cleared before the packet is transmitted.

20. The method of claim 19 further comprising the step of performing a check to ensure that a lock bit in a lock table has been cleared before the packet is transmitted, if the lock bit in the unicast packet buffer has not been cleared.

21. The method of claim 4 further comprising the step of tagging the packet for source learning if the source address is not associated with the first port.

22. The method of claim 21 further comprising the step of associating the destination address with the second port if the destination address is not associated with the second port.

23. The method of claim 22 further comprising, if the flow state is "last multicast packet," the steps of defining the packet for flooding, setting a lock bit in a multicast packet buffer, and changing the flow state to "first unicast packet."

24. The method of claim 23 further comprising the step of specifying the port, the priority, and the bandwidth for the packet.

25. The method of claim 24 further comprising the step of setting a lock bit in a lock table if the lock bit is set in the multicast packet buffer.

26. The method of claim 25, wherein the step of storing the packet comprises the step of storing the packet in a multicast queue until the specified bandwidth is available at the specified priority for the specified port.

27. The method of claim 26 further comprising the step of clearing the lock bit in the lock table when the packet is transmitted.

28. The method of claim 27 further comprising the step of sending the packet to a source learning element if the source address is not associated with the first port.

29. The method of claim 28 further comprising the step of source learning in which the source address is associated with the first port and association information is stored for inquiries or transmittal.

30. The method of claim 22 further comprising the step of setting a lock bit in a unicast packet buffer if the flow state is "first unicast packet."

31. The method of claim 30 further comprising the step of specifying the port, the priority, and the bandwidth for the packet.

32. The method of claim 31 wherein the step of storing the packet comprises the step of storing the packet in a unicast queue until the specified bandwidth is available at the specified priority for the specified port.

33. The method of claim 32 further comprising the step of performing a check to ensure that the lock bit in the unicast packet buffer has been cleared before the packet is transmitted.

34. The method of claim 33 further comprising, if the lock bit in the unicast packet buffer has not been cleared, the step of performing a check to ensure that a lock bit in a lock table has been cleared before the packet is transmitted.

35. The method of claim 34 further comprising the step of sending the packet to a source learning element if the source address is not associated with the first port.

36. The method of claim 35 further comprising the step of source learning in which the source address is associated with the first port and association information is stored for inquiries or transmittal.

37. The method of claim 22 further comprising the step of: specifying the port, the priority, and the bandwidth for the packet if the flow state is neither "last multicast packet" nor "first unicast packet."

38. The method of claim 37 wherein the step of storing the packet comprises the step of storing the packet in a unicast queue until the specified bandwidth is available at the specified priority for the specified port.

39. The method of claim 38 further comprising the step of performing a check to ensure that a lock bit in a unicast packet buffer has been cleared before the packet is transmitted.

40. The method of claim 39 further comprising the step of performing a check to ensure that a lock bit in a lock table has been cleared before the packet is transmitted, if the lock bit in the unicast packet buffer has not been cleared.

41. The method of claim 40 further comprising the step of sending the packet to a source learning element if the source address is not associated with the first port.

42. The method of claim 41 further comprising the step of source learning in which the source address is associated with the first port and association information is stored for inquiries or transmittal.

43. The method of claim 4 further comprising, if the destination address is not associated with the second port, the steps of defining the packet for flooding; and notifying a source learning element.

44. The method of claim 43 further comprising the step of specifying the port, the priority, and the bandwidth for the packet.

45. The method of claim 44 further comprising the step of setting a lock bit in a lock table if a lock bit is set in a multicast packet buffer.

46. The method of claim 45 wherein the step of storing the packet comprises the step of storing the packet in a multicast queue until the specified bandwidth is available at the specified priority for the specified port.

47. The method of claim 46 further comprising the step of clearing the lock bit in the lock table when the packet is transmitted.

48. The method of claim 47 further comprising the step of sending the packet to a source learning element if the source address is not associated with the first port.

49. The method of claim 48 further comprising the step of source learning in which the source address is associated with the first port and association information is stored for inquiries or transmittal.

50. The method of claim 43 further comprising the step of searching for the destination address while packets are being flooded.

51. The method of claim 50 further comprising, when the destination address has been found, the steps of: specifying the port, the priority, and the bandwidth for the packet; storing the association between the destination address and the second port; and setting the flow state to "last multicast packet."

52. The method of claim 21 further comprising, if the destination address is not associated with the second port, the steps of: defining the packet for flooding; and notifying a source learning element.

53. The method of claim 52 further comprising the step of specifying the port, the priority, and the bandwidth for the packet.

54. The method of claim 53 further comprising the step of setting a lock bit in a lock table if a lock bit is set in a multicast packet buffer.

55. The method of claim 54 wherein the step of storing the packet comprises the step of storing the packet in a multicast queue until the specified bandwidth is available at the specified priority for the specified port.

56. The method of claim 55 further comprising the step of clearing the lock bit in the lock table when the packet is transmitted.

57. The method of claim 56 further comprising the step of sending the packet to a source learning element if the source address is not associated with the first port.

58. The method of claim 57 further comprising the step of source learning in which the source address is associated with the first port and association information is stored for inquiries or transmittal.

59. The method of claim 52 further comprising the step of searching for the destination address while packets are being flooded.

60. The method of claim 59 further comprising, when the destination address is found, the steps of: specifying the port, the priority, and the bandwidth for the packet; storing the association between the destination address and the second port; and setting the flow state to "last multicast packet."

61. A data communication switch, comprising:

a plurality of switching modules;

a backplane coupled to the switching modules for exchanging packet data originated by and destined to external network devices; and

a control path coupled to the switching modules for exchanging control data originated by and destined to the switching modules wherein the control data includes information regarding associations between external network devices and ports of the data communication switch.

62. The data communication switch of claim 61, wherein the switching module comprises:

an access controller having a port for receiving a packet including a destination address and source address;

a switching controller coupled to the access controller for receiving the packet from the access controller and processing it;

a queue controller coupled to the switching controller for receiving the packet from the switching controller,

wherein the queue controller transmits a first plurality of values to a first element in response to the packet, and receives a source address and a first port association from the first element in response to the first plurality of values, transmits a second plurality of values to a second element in response to the packet, and receives a destination address and a second port association from the second element in response to the second plurality of values, transmits a third plurality of values to a third element in response to the packet, and receives a port, a priority, and a bandwidth from the third element in response to the third plurality of values, and transmits the packet using a unicast queue, the destination address, the port, and the priority.

63. The data communication switch of claim 62, wherein the queue controller transmits a fourth plurality of values to a fourth element if the source address and first

port association are not received in response to the first plurality of values, and wherein the queue controller receives the source address and first port association from the fourth element in response to the fourth plurality of values.

64. The data communication switch of claim 63, wherein the queue controller transmits the packet using a multicast queue and transmits a fifth plurality of values to the fourth element if the destination address and the second port association are not received in response to the second plurality of values, and wherein the queue controller receives the destination address and the second port association from the fourth element in response to the fifth plurality of values, delays the packet flow, and transmits the packet using a unicast queue.

DISTRIBUTED SOURCE LEARNING FOR DATA COMMUNICATION SWITCH

ABSTRACT OF THE DISCLOSURE

A method and apparatus for accomplishing source learning in a data switch of the type having a plurality of switching modules where each supports one or more external network devices and a backplane interconnecting the switching modules. Each switching module has logic resident thereon for performing distributed source learning, including configuring unknown source addresses "seen" in inbound packets and for notifying the other switching modules that such source addresses were "seen" on a port thereof. Address-port associations are thereby configured on the switch using distributed logic, i.e. without intervention by a centralized management entity. In regard to configuring destination addresses--when a destination address is unknown, packets are delivered over a multicast queue until the destination address is found. Once the destination address is found, a method of flow integrity is used to avoid out of order packet delivery when the device transitions from using a multicast flood queue to a unicast queue.

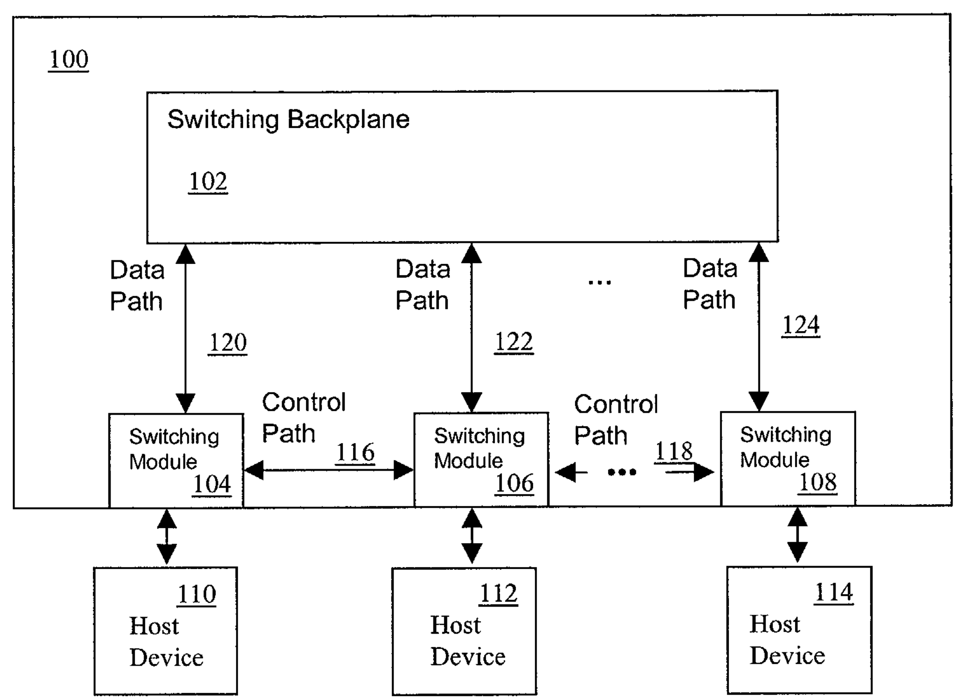
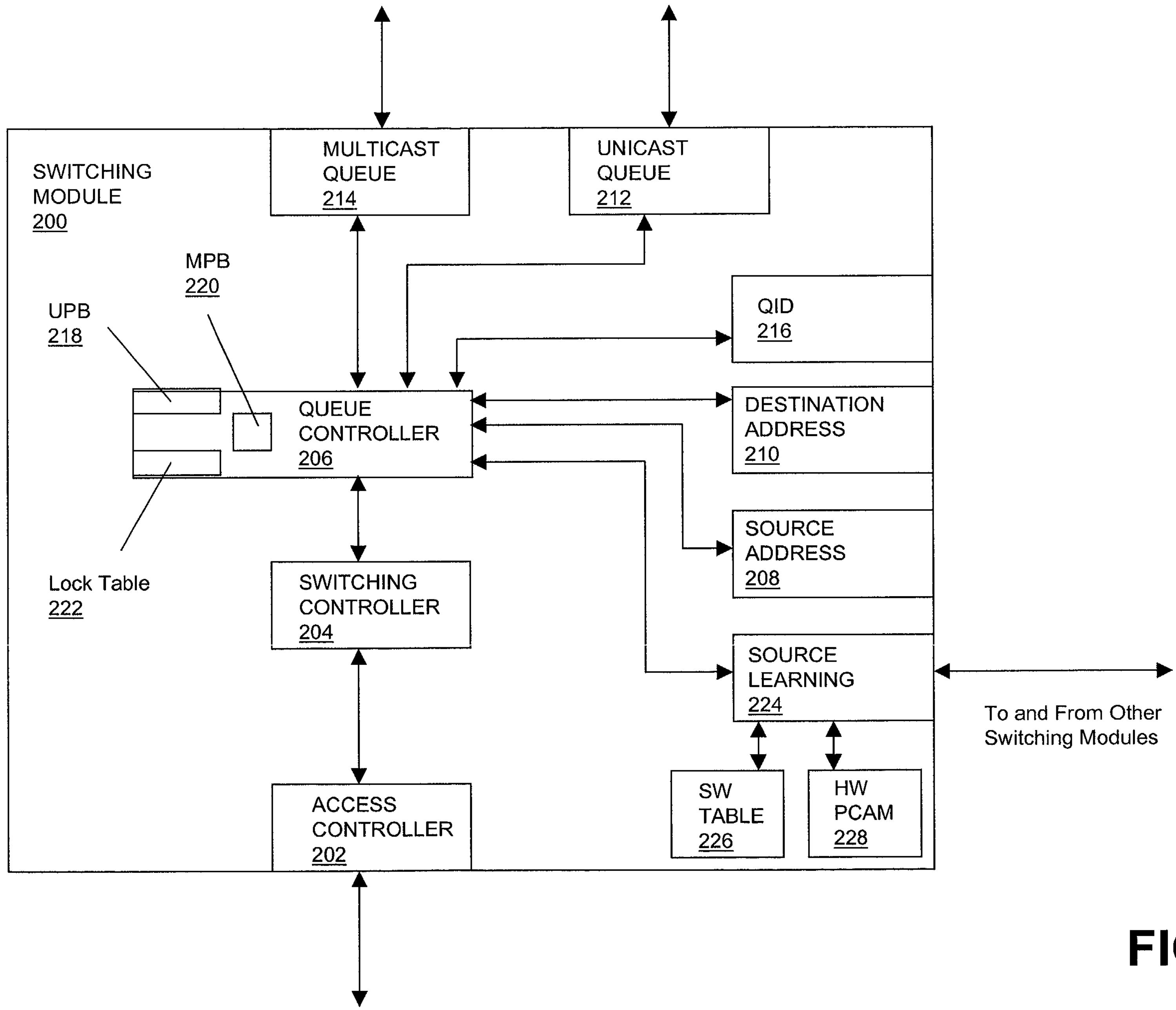

100
Switching Backplane
102
Data Path
120
Data Path
122
...
Data Path
124
Control Path
116
Control Path
118
Switching Module 104
Switching Module 106
Switching Module 108
110 Host Device
112 Host Device
114 Host Device

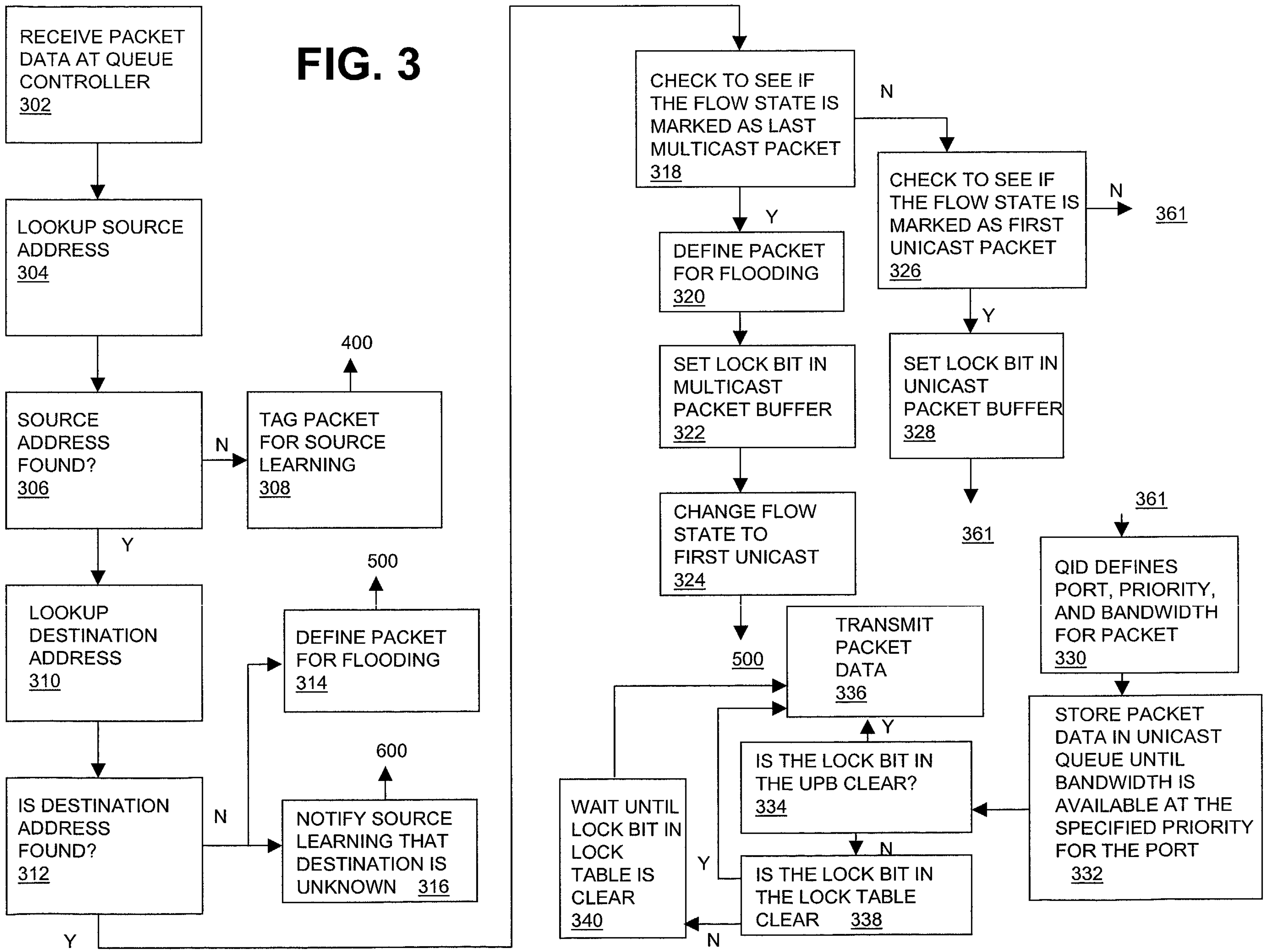
FIG. 1



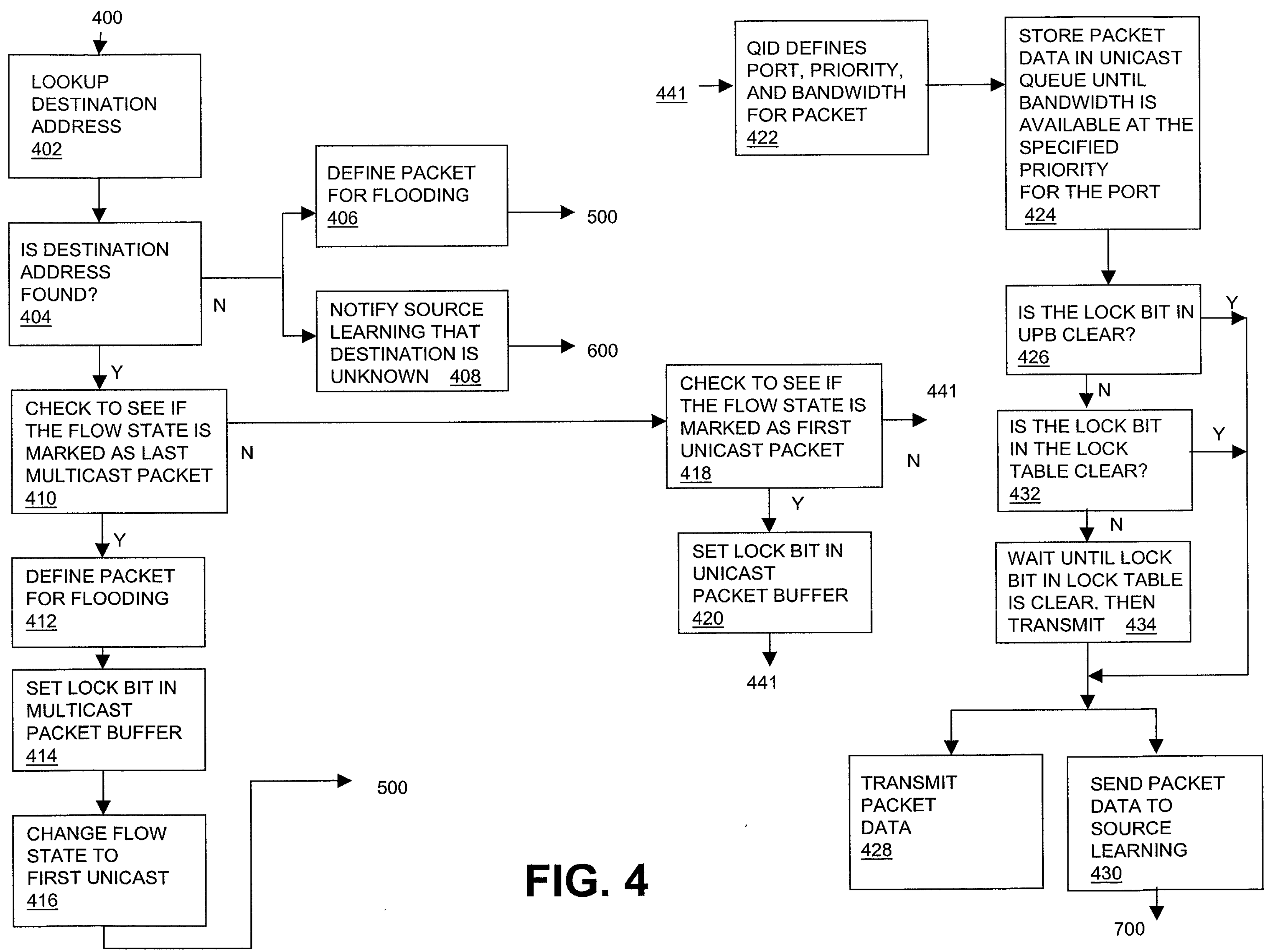
SWITCHING MODULE 200
MULTICAST QUEUE 214
UNICAST QUEUE 212
MPB 220
UPB 218
QID 216
QUEUE CONTROLLER 206
DESTINATION ADDRESS 210
SOURCE ADDRESS 208
Lock Table 222
SWITCHING CONTROLLER 204
SOURCE LEARNING 224
To and From Other Switching Modules
SW TABLE 226
HW PCAM 228
ACCESS CONTROLLER 202

FIG. 2

FIG. 3



RECEIVE PACKET DATA AT QUEUE CONTROLLER 302
LOOKUP SOURCE ADDRESS 304
SOURCE ADDRESS FOUND? 306
N
TAG PACKET FOR SOURCE LEARNING 308
400
Y
LOOKUP DESTINATION ADDRESS 310
IS DESTINATION ADDRESS FOUND? 312
N
DEFINE PACKET FOR FLOODING 314
500
NOTIFY SOURCE LEARNING THAT DESTINATION IS UNKNOWN 316
600
Y
CHECK TO SEE IF THE FLOW STATE IS MARKED AS LAST MULTICAST PACKET 318
N
Y
DEFINE PACKET FOR FLOODING 320
SET LOCK BIT IN MULTICAST PACKET BUFFER 322
CHANGE FLOW STATE TO FIRST UNICAST 324
500
CHECK TO SEE IF THE FLOW STATE IS MARKED AS FIRST UNICAST PACKET 326
N
361
Y
SET LOCK BIT IN UNICAST PACKET BUFFER 328
361
361
QID DEFINES PORT, PRIORITY, AND BANDWIDTH FOR PACKET 330
STORE PACKET DATA IN UNICAST QUEUE UNTIL BANDWIDTH IS AVAILABLE AT THE SPECIFIED PRIORITY FOR THE PORT 332
TRANSMIT PACKET DATA 336
Y
IS THE LOCK BIT IN THE UPB CLEAR? 334
N
IS THE LOCK BIT IN THE LOCK TABLE CLEAR 338
Y
N
WAIT UNTIL LOCK BIT IN LOCK TABLE IS CLEAR 340



400
LOOKUP DESTINATION ADDRESS 402
IS DESTINATION ADDRESS FOUND? 404
N
Y
DEFINE PACKET FOR FLOODING 406
500
NOTIFY SOURCE LEARNING THAT DESTINATION IS UNKNOWN 408
600
CHECK TO SEE IF THE FLOW STATE IS MARKED AS LAST MULTICAST PACKET 410
N
Y
DEFINE PACKET FOR FLOODING 412
SET LOCK BIT IN MULTICAST PACKET BUFFER 414
CHANGE FLOW STATE TO FIRST UNICAST 416
500
CHECK TO SEE IF THE FLOW STATE IS MARKED AS FIRST UNICAST PACKET 418
441
N
Y
SET LOCK BIT IN UNICAST PACKET BUFFER 420
441
441
QID DEFINES PORT, PRIORITY, AND BANDWIDTH FOR PACKET 422
STORE PACKET DATA IN UNICAST QUEUE UNTIL BANDWIDTH IS AVAILABLE AT THE SPECIFIED PRIORITY FOR THE PORT 424
IS THE LOCK BIT IN UPB CLEAR? 426
Y
N
IS THE LOCK BIT IN THE LOCK TABLE CLEAR? 432
Y
N
WAIT UNTIL LOCK BIT IN LOCK TABLE IS CLEAR. THEN TRANSMIT 434
TRANSMIT PACKET DATA 428
SEND PACKET DATA TO SOURCE LEARNING 430
700

FIG. 4

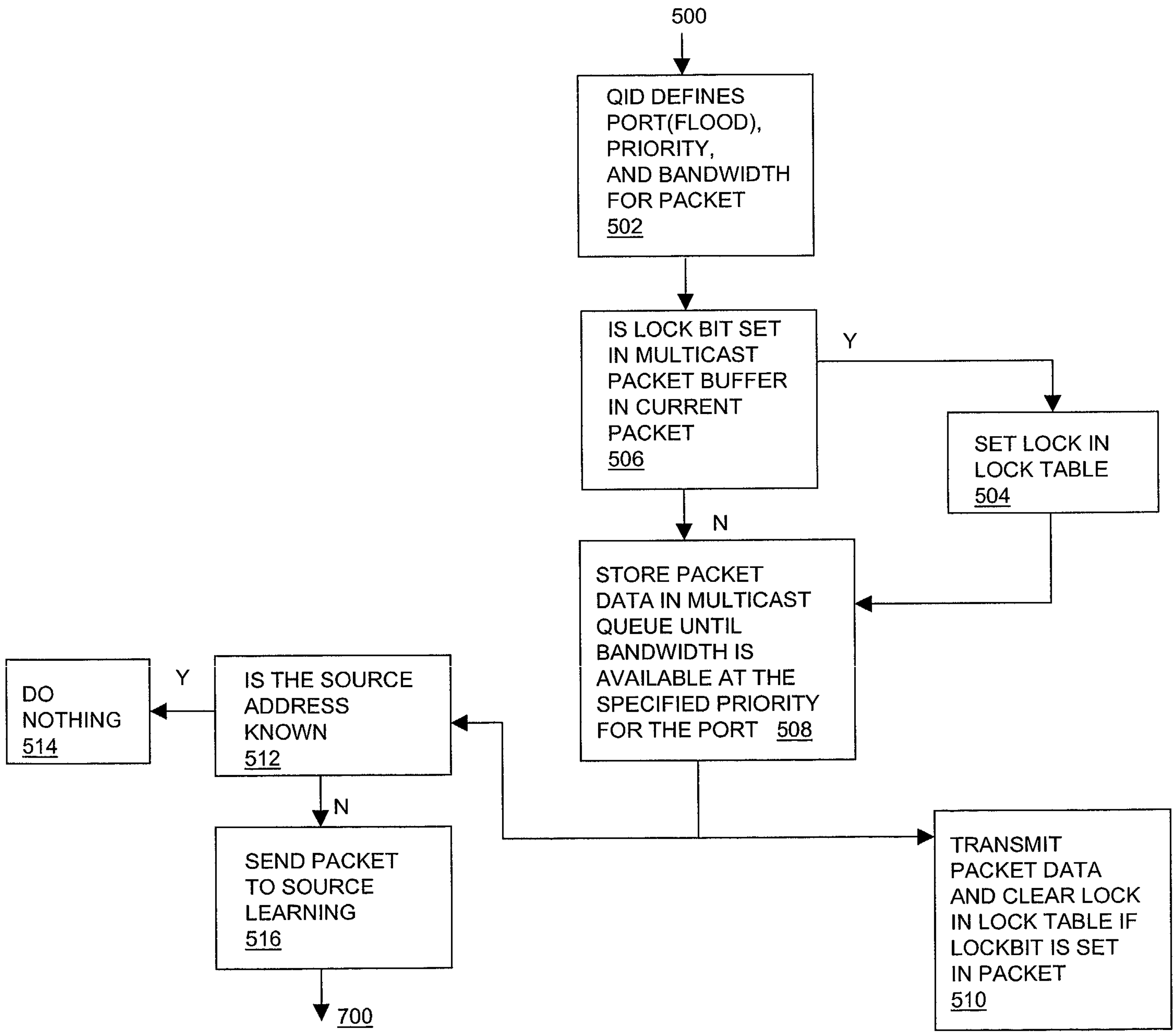

500
QID DEFINES PORT(FLOOD), PRIORITY, AND BANDWIDTH FOR PACKET 502
IS LOCK BIT SET IN MULTICAST PACKET BUFFER IN CURRENT PACKET 506
Y
SET LOCK IN LOCK TABLE 504
N
STORE PACKET DATA IN MULTICAST QUEUE UNTIL BANDWIDTH IS AVAILABLE AT THE SPECIFIED PRIORITY FOR THE PORT 508
TRANSMIT PACKET DATA AND CLEAR LOCK IN LOCK TABLE IF LOCKBIT IS SET IN PACKET 510
IS THE SOURCE ADDRESS KNOWN 512
Y
DO NOTHING 514
N
SEND PACKET TO SOURCE LEARNING 516
700

FIG. 5

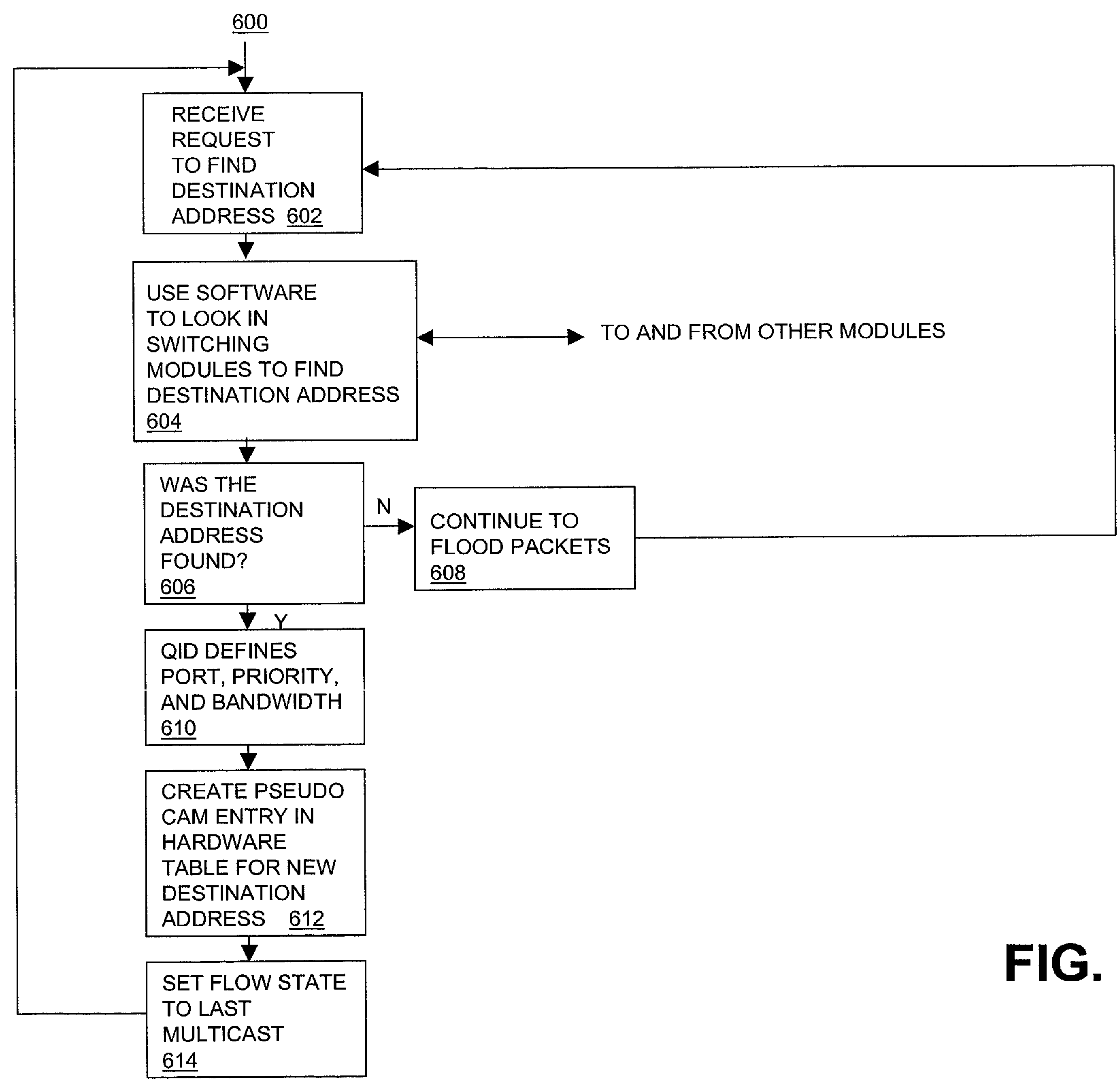
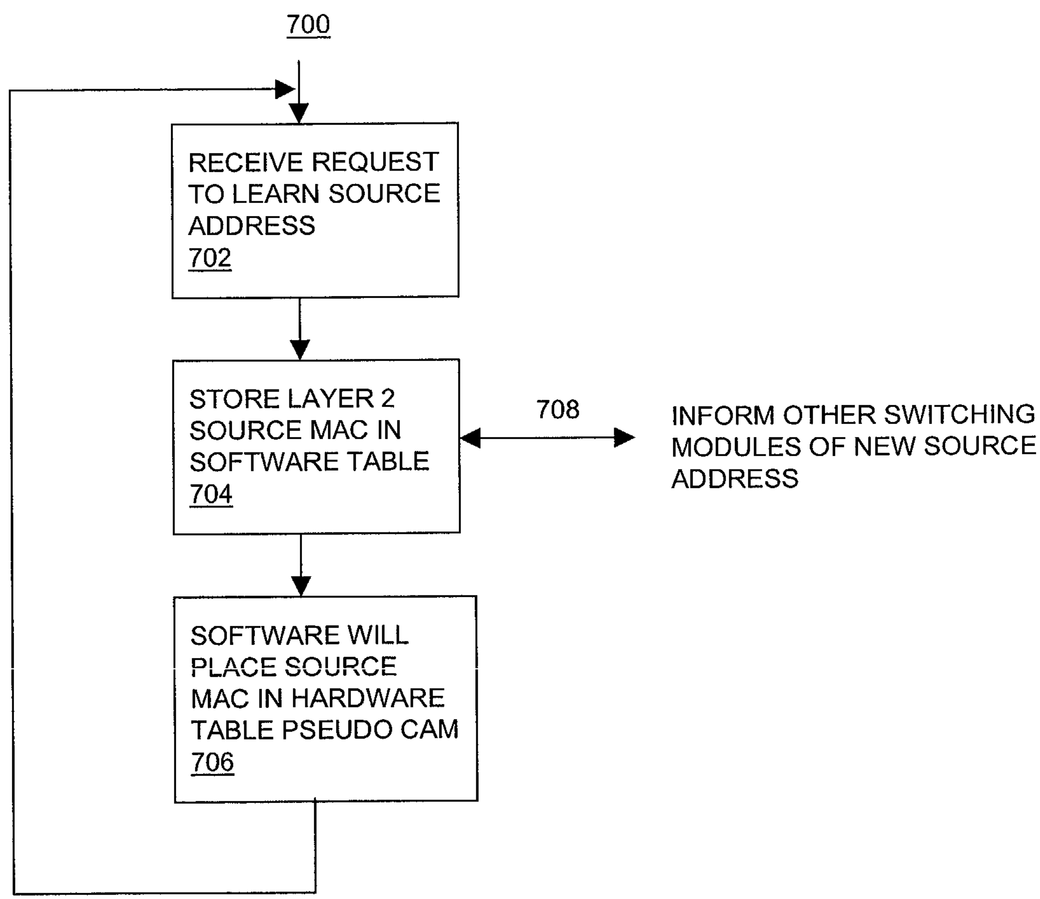

600
RECEIVE REQUEST TO FIND DESTINATION ADDRESS 602
USE SOFTWARE TO LOOK IN SWITCHING MODULES TO FIND DESTINATION ADDRESS 604
TO AND FROM OTHER MODULES
WAS THE DESTINATION ADDRESS FOUND? 606
N
CONTINUE TO FLOOD PACKETS 608
Y
QID DEFINES PORT, PRIORITY, AND BANDWIDTH 610
CREATE PSEUDO CAM ENTRY IN HARDWARE TABLE FOR NEW DESTINATION ADDRESS 612
SET FLOW STATE TO LAST MULTICAST 614

FIG. 6



700
RECEIVE REQUEST TO LEARN SOURCE ADDRESS 702
STORE LAYER 2 SOURCE MAC IN SOFTWARE TABLE 704
708
INFORM OTHER SWITCHING MODULES OF NEW SOURCE ADDRESS
SOFTWARE WILL PLACE SOURCE MAC IN HARDWARE TABLE PSEUDO CAM 706

FIG. 7

DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATIONS PATENT

Docket No. : 39981/DAC/X2

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled DISTRIBUTED SOURCE LEARNING FOR DATA COMMUNICATION SWITCH, the specification of which is attached hereto unless the following is checked:

__ was filed on __ as United States Application Number or PCT International Application Number __ and was amended on ___ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR § 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. § 119(a)-(d) or § 365(b) of the foreign application(s) for patent or inventor's certificate, or § 365(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below, any foreign application for patent or inventor's certificate, or PCT International application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application(s)

Application Number	Country	Filing Date (day/month/year)	Priority Claimed

I hereby claim the benefit under 35 U.S.C. § 119(e) of any United States provisional application(s) listed below.

Application Number	Filing Date

I hereby claim the benefit under 35 U.S.C. § 120 of any United States application(s), or any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. § 112, I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR § 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application:

Application Number	Filing Date	Patented/Pending/Abandoned

POWER OF ATTORNEY: I hereby appoint Scot A. Reader (39,002) of Alcatel Internetworking, Inc. and the following attorneys and agents of the law firm CHRISTIE, PARKER & HALE, LLP to prosecute this application and any international application under the Patent Cooperation Treaty based on it and to transact all business in the U.S. Patent and Trademark Office connected with either of them in accordance with instructions from the assignee of the entire interest in this application; or from the first or sole inventor named below in the event the

application is not assigned; or from __ in the event the power granted herein is for an application filed on behalf of a foreign attorney or agent.

R. W. Johnston	(17,968)	Gregory S. Lampert	(35,581)	Raymond R. Tabandeh	(43,945)
D. Bruce Prout	(20,958)	Grant T. Langton	(39,739)	Phuong-Quan Hoang	(41,839)
Hayden A. Carney	(22,653)	Constantine Marantidis	(39,759)	Cynthia A. Bonner	(44,548)
Richard J. Ward, Jr.	(24,187)	Marilyn R. Khorsandi	(45,744)	Jun-Young E. Jeon	(43,693)
Russell R. Palmer, Jr.	(22,994)	Daniel R. Kimbell	(34,849)	Marc A. Karish	(44,816)
LeRoy T. Rahn	(20,356)	Craig A. Gelfound	(41,032)	John F. O'Rourke	(38,985)
Richard D. Seibel	(22,134)	Syed A. Hasan	(41,057)	Richard J. Paciulan	(28,248)
Walter G. Maxwell	(25,355)	Kathleen M. Olster	(42,052)	Josephine E. Chang	(46,083)
William P. Christie	(29,371)	Daniel M. Cavanagh	(41,661)	Frank L. Cire	(42,419)
David A. Dillard	(30,831)	Molly A. Holman	(40,022)	Harold E. Wurst	(22,183)
Thomas J. Daly	(32,213)	Lucinda G. Auciello	(42,270)	Robert A. Green	(28,301)
Vincent G. Gioia	(19,959)	Norman E. Carte	(30,455)	Derrick W. Reed	(40,138)
Edward R. Schwartz	(31,135)	Joel A. Kauth	(41,886)	Stephen D. Burbach	(40,285)
John D. Carpenter	(34,133)	Patrick Y. Ikehara	(42,681)	David B. Sandelands, Jr.	(46,023)
David A. Plumley	(37,208)	Mark Garscia	(31,953)	Heidi L. Eisenhut	(P-46,812)
Wesley W. Monroe	(39,778)	Gary J. Nelson	(44,257)		

The authority under this Power of Attorney of each person named above of the law firm shall automatically terminate and be revoked upon such person ceasing to be a member or associate of or of counsel to that law firm.

DIRECT TELEPHONE CALLS TO : Art Hasan, 626/795-9900

SEND CORRESPONDENCE TO : CHRISTIE, PARKER & HALE, LLP
P.O. Box 7068, Pasadena, CA 91109-7068

I declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of sole or first joint inventor	Inventor's signature	Date
Rex A. Hill		
Residence and Post Office Address		Citizenship
14360 Janal Way, San Diego, California 92129		USA

JEJ PAS268099 1-*-8/16/00 11 40 AM